A Closed-form Focus Profile Model for Conventional Digital Cameras

Supplemental material

S. Pertuz, M.A. Garcia, D. Puig, H. Arguello

1 Sampling and quantization

This section describes in more detail the steps that take place in the digitization of an image: sampling and quantization.

1.1 Sampling

In the process of sampling, a continuous 2D irradiance field $I_c(x_c, y_c)$ is mapped to a discrete 2D image, $I(x, y)$. In this document, x and y are discrete spatial variables, whereas the sub-index c has been used to indicate continuity in the space domain. With this notation, the sampling of the continuous signal $I_c(x_c, y_c)$ into its discrete counterpart can be described as [34]:

$$I_c(x_c, y_c) \rightarrow I(x, y) = I_c(mS_x, nS_y), \tag{1}$$

where S_x and S_y are the sample spacings, in the x and y directions, respectively; $m = \{1, 2, \cdots, M\}$ and $n = \{1, 2, \cdots, N\}$ are the discrete image indices, assuming that the discrete 2D image has $M \times N$ samples.

The sampling rule in (1) is an ideal representation that takes samples of the continuous image irradiance on infinitesimal 2D spots at coordinates (mS_x, nS_y). However, in practice, the sensing device relies on processing a small area of finite size in order to estimate the energy of the irradiance corresponding to a single location. Specifically, each pixel consists of a photosensor aimed at measuring the image irradiance by integrating it both in the spatial and temporal domains. For illustration purposes, let us consider the simplest case of an individual imaging sensor consisting of a matrix of squared photoreceptors (see Fig. 1). In this case, the image irradiance is integrated over the area of each individual photoreceptor. Thus, the sample at coordinates (x, y) corresponds to [12]:

$$I(x, y) = G\beta\lambda\frac{\tau}{hc}\int\int_{\Omega(x_c, y_c)} I(x_c, y_c)\, dx_c\, dy_c, \tag{2}$$

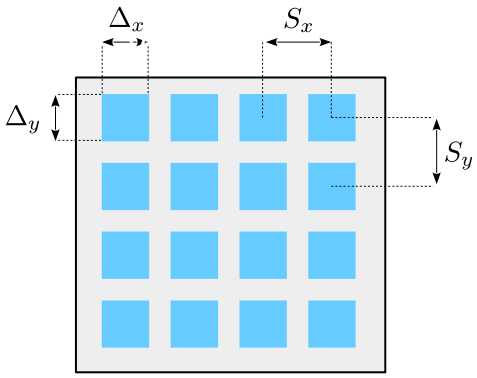

Figure 1: Image sampling. $\Delta_x \times \Delta_y$ is the detector size and S_x and S_y are the sample spacings in the horizontal and vertical directions, respectively.

where $\Omega(x_c, y_c)$ is a neighborhood that covers the $\Delta_x \times \Delta_y$ pixel area so that $\Omega(x_c, y_c) = \{(x, y)|x \in [x_c - 0.5\Delta_x, x_c + 0.5\Delta_x] \wedge y \in [y_c - 0.5\Delta_y, y_c + 0.5\Delta_y]\}$, G is the sensor gain constant, β is the quantum efficiency of the photosensor (the conversion rate between incoming photons and unit charges), λ is the wavelength of the incoming light, τ is the integration time (exposure) and h and c are the Plank's constant and the speed of light, respectively [1].

The effect of integrating over a finite spot in (2) combined with the sampling of (1), can be interpreted as applying an averaging filter and the multiplication with an impulse train [9]:

$$I(x, y) = I(x_c, y_c) * A(x_c, y_c) \sum_{m=1}^{M} \sum_{n=1}^{N} \delta(x_c - nS_x, y_c - mS_y), \tag{3}$$

where $A(x_c, y_c)$ is a bi-dimensional squared window of size $\Delta_x \times \Delta_y$ and $\delta(x_c, y_c)$ is the impulse function.

By simple inspection of the convolution in (3), it is clear that the image irradiance is low-pass filtered by the sampling spot. Thus, the sampling spot may be utilized to perform pre-sampling filtering for avoiding aliasing. Aliasing is an undesired effect that can appear when representing a continuous domain function with a discrete number of samples. Specifically, according to the Nyquist-Shannon sampling theorem, the continuous domain irradiance, I_c, must be band-limited in order to avoid *aliasing* [10]:

$$\xi_c < \frac{1}{2S_x}, \ \eta_c < \frac{1}{2S_y}, \tag{4}$$

[1] Strictly, the quantum efficiency is a function of the wavelength: $\beta = f(\lambda)$. Therefore, in polychromatic illumination, (2) involves an integral as a function of λ [12]. For simplicity, this step has been omitted since it is not essential for the subsequent discussion.

where ξ_c and η_c are the cut-off frequencies of I_c in the x and y directions, respectively.

Bearing the previous discussion in mind, the selection of the sampling steps, S_x and S_y, as well as the dimension of the sampling spot, $\Delta_x \times \Delta_y$ is an important design criterion that determines the final resolution of the imaging system [25]. It is important to remark that, even in perfect focus, the imaged target is blurred due to the effect of diffraction. This diffraction blurring can be interpreted as a band-limiting effect of the system's optics. As a result, in order to optimize the acquisition process, the sensor resolution must be designed to be as close as possible to the resolution of the system's optics at the diffraction limit [5].

1.2 Quantization

During sampling, photodetectors work as transducers that convert the image irradiance into electric charge. In conventional cameras, the output of the detector is then amplified and converted to a discrete value. For instance, in 8-bit monochromatic images (gray-scale images), each pixel is assigned a value between 0 and 255 according to the measured irradiance, being 255 the maximum allowed brightness and 0 the minimum one. Quantization is the process of scaling and discretizing the measured values. Both, photodetector measurement and quantization imply the addition of different noise types to the ideal sensed image. For instance, a CCD camera has several primary noise sources, such as *fixed pattern noise*, *dark current noise*, *shot noise*, *amplifier noise* and *quantization noise* [11], which can be grouped into both irradiance-dependent and irradiance-independent sources. In that way, a noisy image I_n can be modeled as [17]:

$$I_n = g(I + n_s + n_c) + n_q, \tag{5}$$

where I is the original image, $g(\cdot)$ is the camera response function (CRF), n_s is the irradiance-dependent noise component, n_c is the independent noise, and n_q is the additional quantization and amplification noise.

2 The Gaussian PSF

When the focus defect is considerable and diffraction effects are neglected, the geometrical optics approximation of the PSF of a defocused system corresponds to the pillbox function. Alternatively, some researchers have suggested using a 2D Gaussian in order to take into account the effects of polychromatic illumination, lens aberrations and other defects. At this point, it is worth asking if the Gaussian approximation is indeed a suitable model for a defocused system.

A thorough review of the literature reveals that the assumption of a Gaussian blur model has widely been accepted and exploited for different applications. One of its main advantages is its mathematical tractability derived from the properties of the Gaussian function. For instance, the Gaussian PSF has successfully been exploited in computer vision for depth retrieval through SFD

[24, 30], for proposing new focus measure operators [37], for computing the all-in-focus image in focus stacking [15, 1] and for image restoration and deblurring [6, 26, 22, 16, 7, 23]. The Gaussian PSF has also been exploited for assessing the effect of defocus blur in human depth perception [19]. Note, however, that some authors have opted for more general models when estimating the PSF of optical systems [31, 36, 14, 8, 13] or for accurately assessing human visual acuity [32, 35].

Previous researchers have shown that the geometrical approximation is similar to the exact, diffracted PSF for large amounts of defocus [28]. Notwithstanding, near the diffraction limit, the Gaussian PSF behaves as an all-pass filter, which neglects the limitations imposed by diffraction. A key aspect of the model proposed in the paper, is that the joint effect of the sampling function and diffraction can be replaced by an ideal low-pass filter. As a result, one can take advantage of the limited resolution of real imaging systems in order to derive an approximated model even at perfect focus. Note, however, that this approximation is not suitable for modeling the PSF of a defocused optical system and is, therefore, not suitable for some applications, such as image deblurring based on direct deconvolution.

3 Experiments in Shape-from-focus

This section provides experiments on the application of the proposed focus profile model for depth estimation using *shape-from-focus*.

Shape-from-focus can be divided into three main stages: focus measure, peak detection and post-processing. Once a focus sequence has been captured by changing focus, the *focus measure* stage consists of the application of a focus measure operator in order to compute the focus profile for each pixel of an imaged scene. Traditionally, *peak detection* has been performed by fitting a model, $\tilde{\varphi}$, to the measured focus profile, φ. The depth of the corresponding pixels is estimated by simply finding the maximum of the interpolated focus profile model. The obtained depth-map can be further improved by means of specific *post-processing* techniques that take into account the continuous nature of the imaged scene [29]. A wide variety of post-processing techniques have been proposed, including surface fitting and optimization by neural networks [4], and dynamic programming [2, 3], among many others (e.g., [27, 20, 18]).

Formally, the aim of SFF is to generate a depth-map, $z(x, y)$, whit the estimated depth at each (x, y) coordinate. For comparison purposes, the estimated depth has been computed using the Gaussian profile [21], the Laplacian-Cauchy profile [33, 20] and the proposed focus profile model. Since the aim of this experiment is to compare the different focus profile models, the presented results correspond to raw depth-maps without any post-processing stage.

In order to simplify the construction and measurement of the ground truth, the captured scenes mostly consisted of planar objects placed at different positions from the camera. For each scene, a focus sequence of 100 frames was acquired using a Sony SNC-RZ50P camera. The depth-map was then estimated

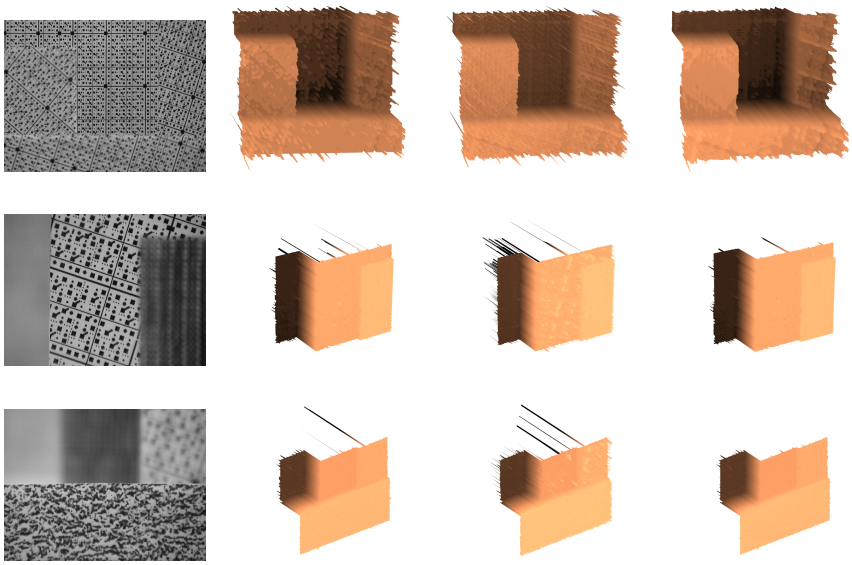

Figure 2: Performance of shape-from-focus using different focus profile models. From left to right: one frame of the focus sequence, Gaussian model, Laplacian-Cauchy model, and proposed model.

Table 1: Performance measures of SFF with different focus profile models: Gaussian [21], GP, Laplacian-Cauchy [20, 33], LC, and Proposed model, FP.

Method	RMSE	AE (%)	SNR (dB)	C
GP	0.108	1.52	35.9	0.972
LC	0.178	2.51	33.4	0.932
FP	**0.095**	**1.41**	**36.6**	**0.975**

by using SFF as described previously.

Five sequences have been used in the experiments and the depth-maps generated with different focus profile models have been compared. In order to measure the quality of the obtained depth-maps, the RMSE, the correlation (C), the signal-to-noise ratio (SNR), and the mean absolute error (AE) have been computed between the reconstructed depth-map, $z(x, y)$, and the ground truth, $g_T(x, y)$. Fig. 2 shows one frame of the focus sequence and the obtained depth-maps for three of the sequences used in the experiments. The mean results are summarized in Table 1. As shown in this table, the best performance of SFF is obtained when the proposed focus profile model is applied.

References

[1] F. Aguet, D. Van De Ville, and M. Unser. Model-based 2.5-D deconvolution for extended depth of field in brightfield microscopy. *IEEE Trans. Image Process.*, 17(7):1144–1153, 2008.

[2] M. B. Ahmad and Tae Sun Choi. A heuristic approach for finding best focused shape. *IEEE Trans. Circuits Syst. Video Technol.*, 15(4):566–574, apr. 2005.

[3] Muhammad Bilal Ahmad and Tae Sun Choi. Shape from focus using optimization technique. In *proc. IEEE Int. Conference on Acoustics, Speech and Signal Processing*, volume 2, pages II –II, may. 2006.

[4] Muhammad Asif and Tae Sun Choi. Shape from focus using multilayer feedforward neural networks. *IEEE Trans. Image Process.*, 10(11):1670–1675, nov. 2001.

[5] Michael Bass, editor. *Handbood of Optics*, volume 1. OSA, 3rd edition, 2010. Geometrical and physical optics, polarized light, components and instruments.

[6] Gang Cao, Yao Zhao, and Rongrong Ni. Edge-based blur metric for tamper detection. *Journal of information Hiding and Multimedia Signal Processing*, 1(1):20–27, 2010.

[7] Shenyong Chen and Y. F. Li. Finding optima focusing distance and edge blur distribution for weakly calibrated 3d vision. *IEEE Trans. Ind. Inf.*, 9(3):1680–1687, 2013.

[8] Rob Fergus, Barun Singh, Aaron Hertzmann, Sam T. Roweis, and William T. Freeman. Removing camera shake from a single photograph. *ACM Transactions on Graphics*, 25:787–794, 2006.

[9] Robert Fiete. *Modeling the Imaging Chain of Digital Cameras*. SPIE press, 2010.

[10] Joseph W. Goodman. *Introduction to Fourier optics*. McGraw-Hill, 2nd edition, 1996.

[11] G. Healey and R. Kondepudy. Radiometric CCD camera calibration and noise estimation. *IEEE Trans. Pattern Anal. Mach. Intell.*, 16(3):267–276, 1994.

[12] Bernard Jähne. *Practical Handbook on Image Processing for Scientific and Thechnical Applications*. CRC Press, 2nd ed. edition, 2004.

[13] Hui Ji and Kang Wang. Robust image deblurring with an inaccurate blur kernel. *IEEE Transactions on Image Processing*, 21(4):1624 –1634, april 2012.

[14] Neel Joshi, Richard Szeliski, and David Kriegman. PSF estimation using sharp edge prediction. In *proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2008.

[15] K. Kodama, H. Mo, and A. Kubota. Simple and fast all-in-focus image reconstruction based on three-dimensional/two-dimensional transform and filtering. In *proc. IEEE International Conference on Acoustics, Speech and Signal Processing*, volume 1, pages 769–772, apr 2007.

[16] Yang Chin Lai. PSO-based estimation for gaussian blur in blind image deconvolution problem. In *proc. IEEE International Conference on Fuzzy Systems*, pages 1143–1148, 2011.

[17] Ce Liu, R. Szeliski, Sing Bing Kang, C.L. Zitnick, and W.T. Freeman. Automatic estimation and removal of noise from a single image. *IEEE Trans. Pattern Anal. Mach. Intell.*, 30(2):299 –314, feb. 2008.

[18] Muhammad Tariq Mahmood and Tae Sun Choi. Nonlinear approach for enhancement of image focus volume in shape from focus. *IEEE Trans. Image Process.*, 21(5):2866–2873, may. 2012.

[19] George Mather and David R. R. Smith. Blur discrimination and its relation to blur-mediated depth perception. *Perception*, 31:1211–1219, 2002.

[20] M. Muhammad and Tae-Sun Choi. Sampling for shape from focus in optical microscopy. *IEEE Trans. Pattern Anal. Mach. Intell.*, 34(3):564 –573, mar. 2012.

[21] S. K Nayar and Y. Nakagawa. Shape from focus. *IEEE Trans. Pattern Anal. Mach. Intell.*, 16(8):824–831, aug. 1994.

[22] F. Orieux, J. F. Giovanelli, and T. Rodet. Deconvolution with gaussian blur parameter and hyperparameters estimation. In *proc. IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 1350–1353, 2010.

[23] C. Paramanand and A. N. Rajagopalan. Depth from motion and optical blur with an unscented Kalman filter. *IEEE Trans. Image Process.*, 21:2798 – 2811, 2012.

[24] Alex Pentland. A new sense for depth of field. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 9(4):523–531, July 1987.

[25] William K. Pratt. *Digital Image processing: PISK scientific inside.* John Willey & Sons, 4th edition, 2007.

[26] Feng Qing Qin. Blind image surper-resolution reconstruction based on PSF estimation. In *proc. International Conference on Information and Automation*, pages 1200–1203, 2010.

[27] Seong-O Shim and Tae-Sun Choi. A novel iterative shape from focus algorithm based on combinatorial optimization. *Pattern Recognit.*, 43(10):3338 – 3347, 2010.

[28] Per A. Stokseth. Properties of a defocused optica-system. *Journal of the Optical Society of America*, 59:1314–1321, 1969.

[29] M. Subbarao and T.S. Choi. Accurate recovery of three-dimensional shape from image focus. *IEEE Trans. Pattern Anal. Mach. Intell.*, 17(3):266–274, mar. 1995.

[30] Muralidhara Subbarao and Gopal Surya. Depth from defocus: a spatial domain approach. *International Journal of Computer Vision*, 13:271–294, 1994.

[31] Richard Szeliski. *Computer vision: algorithms and applications.* Springer, 2011.

[32] Larry N. Thibos. Retinal image quality for virtual eyes generated by a statistical model of ocular wavefront aberrations. *Ophtalmic and physiological optics*, 29:288–291, 2009.

[33] Dong Chen Tsai and Homer H. Chen. Reciprocal focus profile. *IEEE Trans. Image Process.*, 21(2):459 –468, feb. 2012.

[34] David George Voelz. *Computational Fourier Optics.* SPIE Press., 2010.

[35] Andrew B. Watson and Albert J. Ahumada. Predicting the visual acuity from wavefront aberrations. *Journal of Vision*, 8(4):1–19, 2008.

[36] Tom L. Williams. *The optical transfer function of imaging systems.* CRC Press, 1999.

[37] S. Yousefi, M. Rahman, and N. Kehtarnavaz. A new auto-focus sharpness function for digital and smart-phone cameras. *IEEE Trans. Consum. Electron.*, 57(3):1003 –1009, aug. 2011.